Submission in Response to NSF CI 2030 Request for Information

DATE AND TIME: 2017-04-05 17:48:26
REFERENCE NO: 304

Author Names & Affiliations

- Christopher Romsos - Oregon State University
- Katie Watkins-Brandt - Oregon State University
- Jasmine Nahorniak - Oregon State University

Contact Email Address (for NSF use only)

(Hidden)

Research Domain, discipline, and sub-discipline

Data Systems, Geology, Geophysics, Seafloor mapping, & Geographic Information Systems

Title of Submission

The Regional Class Research Vessel Program’s response to NSF CI 2030

Abstract (maximum ~200 words).

At-sea scientific research cruises are resource limited both in available berths for scientific personnel and as limited duration events in an environment untethered to the wired or mobile (information networks, social networks, or other) networks available onshore. The Regional Class Research Vessel (RCRV) Program has designed and intends to implement an efficient cyberinfrastructure solution to help mitigate these limitations and to enhance our observational, experimental, and analytical capabilities at sea. Specifically, the RCRV solution leverages extra-vessel human and computational resources, facilitates an efficient use of science time at-sea, and maximizes data and information throughput on available communications channels. We believe that the RCRV model for “datapresence” is an example of a holistic view of future needs for innovative cyberinfrastructure to advance the Nation’s ocean research enterprise.

In 2015 we undertook an internal review of data- and telepresence capabilities in the US research fleet, asking some of the very same questions posed in the DCL. Our responses to the DCL questions have been gleaned from this analysis and from our own experiences in developing a comprehensive plan for RCRV datapresence capabilities. In addition, we support the conclusions drawn by UNOLS SatNag as outlined in their separate DCL response.

Question 1 Research Challenge(s) (maximum ~1200 words): Describe current or emerging science or engineering research challenge(s), providing context in terms of recent research activities and standing questions in the field.

A. Remote Participation Challenge
As our colleagues noted in their DCL response from the UNOLS community, remote participation in sea-going oceanographic research is

now a requirement, primarily due to reductions number of sea-going vessels across all vessel classes. Remote participation is enabled through high-availability and high-bandwidth communication channels, the lack of which are limiting the operational effectiveness of current vessels. Enhanced information exchange between shipboard and shoreside operations during oceanographic expeditions can help mitigate challenges where the existing human resources on board are thin and/or at full utilization and can further magnify shipboard science operations and the scientific mission. This need for and challenge from remote participation is not limited to the shipboard science party. Today, many of the ships' integrated machinery and operation systems must remain in regular communication with shoreside ship operations support systems and personnel further taxing the shared communication channels.

B. Bi-Directional Data/Information Exchange Challenge
Augmenting bandwidth across the fleet may go a long way to resolving the remote participation challenge by making teleconferencing and other synchronous person-to-person communications possible. However, more efficient communications protocols, methods, and workflows for leveraging the excess computational capacity onshore are still needed. In other words, there is a challenge before us to enhance machine-to-machine communication between shipboard and shoreside scientific systems vis a vis existing limitations.

This challenge is complicated by the high latency and high packet loss nature of satellite networks as well as human factors related to how different sectors (science, public, commercial) interact with data and information services. The RCRV program has been addressing this challenge by (1) prototyping and testing different models for near real-time sensed data replication from ship to shore and back again, and (2) recognizing and including a suite of interactive tools and services for end-user or consumer data access and visualization as part of an integrated datapresence solution.

C. Adaptability/Scalability Challenge
Seagoing research facilities are multi-disciplinary scientific platforms. With the exception of the R/V Atlantis and R/V Langseth, very few research vessels conduct repeated missions. For this reason research vessels must provide scalable and configurable computing platforms to cope with the changing needs and objectives of each group of visiting scientists. Cloud computing infrastructure and platforms (AWS, Azure, OpenStack, and others) offer excellent solutions for scalability and configurability. However, due to the connectivity bottleneck (challenges A&B) such solutions must be licensed and deployed on-prem (aboard the ship) under a private cloud model before they can offer a managed pool of compute, storage, and networking resources without relying on a ship to shore link to function. As the research science community starts to adopt and develop for cloud computing platforms (such as Google Compute Engine), ocean research platforms will experience increasing demand for onboard computing systems that are compatible.

D. Remote Management & Configuration Control Challenge
Increasing the technical capabilities and complexities in our oceanographic fleet will come with increased support and security challenges for those responsible for managing vessel cyberinfrastructure. Resilience to system failure and security threats requires that we assess and understand risks to our cyberinfrastructure and maintain plans and processes to control how this infrastructure is deployed and configured. Remote management and configuration control must balance security and stability needs with end-user (science) needs for flexibility.

E. Support Challenge
The RCRV model for datapresence, where the vessel operates as a continuous observation platform connected to a user base that extends beyond the traditional limits of the seagoing science party, understands that the need for enhanced accessibility to shipboard sensor and data systems comes with the need for additional support services. This presents an operational challenge for shipboard marine technicians whose duties and responsibilities are already great. Enhanced shipboard sensor and data system complexity must be met with an increase in technical support services for this model to be successful.

End-users, outside the traditional seagoing science party in the continuous observation platform model (e.g. satellite remote sensing validations, climate and weather modeling, environmental monitoring), will demand a consistent and high-quality data product. Providing additional technical oversight and support will ensure real-time quality assessment and quality control of data generated from the RCRV sensor system, as well as providing ease for scientific users to incorporate their own instrumentation and integrate those data streams into the shipboard system for visualization and/or transfer of data to support systems on shore.

F. Workforce Challenge
Compensation for academic support of cyberinfrastructure may never reach a level to be directly competitive with the private sector. For this reason, graduates of traditional computer science training and degree programs don't recruit to our population in high numbers or with regularity. In this case, skilled application, maintenance, and operation of the cyberinfrastructure are the challenges. As we modernize our fleets the need for highly skilled technicians and system administrators will increase.

Question 2 Cyberinfrastructure Needed to Address the Research Challenge(s) (maximum ~1200 words): Describe any limitations or absence of existing cyberinfrastructure, and/or specific technical advancements in cyberinfrastructure (e.g. advanced computing, data infrastructure, software infrastructure, applications, networking, cybersecurity), that must be addressed to accomplish the identified research challenge(s).

A. Remote Participation CI Need
Bandwidth is the primary limiting factor limiting remote participation. Bandwidth expansions for ocean research platforms are not keeping pace with the needs and expectations of users. A tenfold increase in bandwidth provision per ship, for example from 1 Mbps to 10 Mbps, would help mitigate this problem by providing adequate bandwidth to permit researchers to engage in small audience telepresence activities without the need for a special bandwidth expansion event. In addition to the obvious operational costs of increased bandwidth purchases there are capital needs and costs to consider here. As noted in the UNOLS response, modernization of the HiSeasNet earth station at UCSD is an important capital expense to undertake. We agree that time and efficiency savings, and ultimately a higher quality of research products from facility users, would offset increased operational expenses resulting from increased bandwidth provision.

B. Bi-Directional Data/Information Exchange CI Need
With adequate bandwidth allocation in place across the fleet, we can begin to improve upon the limited and asynchronous data and information exchange challenge. One component of the RCRV design process has been to take ownership of this challenge and to design the facility to be ready to operate as a continuous observation platform pushing near real-time observational data collected from resident shipboard sensors to a shoreside data store.
Operating as a real-time or near real-time data collection facility requires that we implement standardized services for data access on top of a shoreside content delivery network in parallel to the traditional pathway for data assembly, data curation, and data archive & distribution offered through the National Center for Environmental Information. The RCRV datapresence model uses well-known and stable data access protocols and services (OPeNDAP, ERDDAP, Web Map Services, etc.) to enable real-time data access. The RCRV model also provides a web interface for users to configure personalized alerts and notifications.

C. Adaptability/Scalability CI Need
Data volume, velocity, and variety are increasing; we must ensure that our shipboard cyberinfrastructure is up to the task of managing and protecting these data. To that end, cloud service providers offer numerous solutions for scalability. Given that cruises are multi-disciplinary events of limited duration, seagoing research facilities could potentially benefit from an on-premise and private cloud implementation (such as open stack or similar). To our knowledge we’ve yet to see this rolled out in production on any US research ship.

The UNOLS DCL response indicates a need for enterprise grade hardware and wider use of a virtualized infrastructure. We’d take that a step further and argue that enterprise and datacenter approaches to scalability, redundancy, and availability (not just enterprise hardware) need to be adopted and implemented across our existing and new facilities. That is, our cyberinfrastructure plans should include support for identifying, evaluating, and adopting current computing practices that keep us aligned with the direction of the industry.

D. Remote Management & Configuration Control CI Need
There are various commercial and open source management and monitoring toolkits available for this purpose. However, proper organizational structure and personnel are required to implement them effectively. The RCRV approach to this challenge has been to provide oversight and configuration control for the vessel class through the activities of the Datapresence Systems Engineer and staff at the RCRV Class Management Office.

E. Support CI Need
The RCRV solution to the technical oversight and support challenge is to increase the level of support for shipboard systems through the addition of a shore based Sensor Technician. The success of these systems will be highly reliant on the support provided. A few current examples of this support model include the highly regarded University of Hawaii Data Acquisition System (UHDAS) Acoustic Doppler Current profiler support group and the Shipboard Automated Meteorological and Oceanographic System (SAMOS) support group at Florida State University. The RCRV datapresence plan develops a similar shoreside support expert position for a “Sensor Technician”. The sensor technician has varied responsibilities including maintaining the condition and calibration of RCRV resident sensors. The sensor technician shall also provide an added level of real-time quality assessment for RCRV underway data. The effectiveness of this quality oversight will rely heavily on the real-time datapresence capabilities of the ship.

F. Workforce CI Need

The CI Workforce need is really a recruitment and retention problem. How do we recruit and retain highly skilled computer scientists to support advanced cyberinfrastructure without competing directly with private industry? One potential solution to the future cyberinfrastructure recruitment workforce need is for the NSF to work with university computer science departments to sponsor specific cyberinfrastructure training programs and/or opportunities for undergraduates and graduates in traditional science programs. A potential solution for retention, that doesn't increase the compensation rate, would be one that ensured long-term employment stability (job security) as offsetting a lower pay rate.

Question 3 Other considerations (maximum ~1200 words, optional): Any other relevant aspects, such as organization, process, learning and workforce development, access, and sustainability, that need to be addressed; or any other issues that NSF should consider.

Cost-Benefit Analysis of Increased Bandwidth:
It would be helpful to conduct a formal bandwidth cost-benefit analysis for oceanographic research ships. Some quantification of both the direct and indirect needs and associated benefits would permit an informed discussion and proper bandwidth allocation for our fleet. With a new generation of Ocean Class ships now online, and a new generation of Regional Class ships on the way, now is the time to characterize the increased communication needs due to "smart" technologies embedded in ships equipment and the emerging needs and uses for science. The cost-benefit analysis should also consider the needs, desires, and attitudes of the scientific community particularly with respect to how a new allocation might enhance/enable the science mission.
Open Source Collaborations:
The current paucity of shipboard data visualization tools, data infrastructure, and software infrastructure, places limits on the scientific users and their missions while at sea. With enhanced sensor and data systems comes a need to visualize data in near real time—a capability that is not widely available on most vessels—to help guide the scientific mission. Together with bidirectional data flow capabilities, this will further enhance the mission emphasizing shore-based participation facilitating quality assessment and quality control of the sensor and data systems. This does however require a sophisticated system for data infrastructure and application delivery putting it out of reach from some operators.
Frameworks for open source development and collaboration should be supported to mitigate this challenge. Furthermore, support for technology transfer from those leading the development and implementation to those who might need or want to adopt new cyberinfrastructure, should be part of the development and collaboration framework.